

# **Détection d'anomalies réelles**

**dans les vidéos de surveillance**

**Hao ZHANG**

# Contexte et motivation de la recherche

## Context:

- Les caméras de surveillance sont de plus en plus utilisées dans les lieux publics afin de renforcer la sécurité publique ;
- Les forces de l'ordre ne sont pas en mesure de maintenir leurs capacités de surveillance à niveau, ce qui entraîne une sous-utilisation importante des caméras de surveillance.

# Contexte et motivation de la recherche

## Motivation :

- Il est urgent de développer des algorithmes intelligents pour la détection automatique des anomalies vidéo afin de réduire les pertes de temps et de main-d'œuvre.

# Limites des méthodes précédentes examinées dans cette étude

## Détection d'événements anormaux spécifiques :

- Par exemple, détection de la violence, détection des accidents de la circulation.
- Application limitée dans la pratique en raison du problème de généralité et de l'impossibilité d'étendre la détection à d'autres événements.

# Limites des méthodes précédentes examinées dans cette étude

## Approches basées sur le codage parcimonieux :

- Ces approches supposent que la partie initiale de la vidéo contient des événements normaux afin de construire un dictionnaire d'événements normaux.
- Ne gère pas bien les changements environnementaux ;
- Produit des taux élevés de fausses alarmes pour différents comportements normaux.

# Méthodologie fournie par l'étude

## Objectif :

- Signaler à temps les écarts par rapport aux schémas d'activité normaux dans les vidéos de surveillance de longue durée ;
- Identifier les fenêtres temporelles pendant lesquelles des anomalies se produisent.



# Méthodologie fournie par l'étude

## Méthodes de recherche :

- Vidéos d'entraînement utilisant des étiquettes faibles ;
- Apprendre les anomalies en utilisant des vidéos normales et anormales ;
- Utilisation de l'approche MIL (Multiple Instance Learning) et du modèle de classement MIL profond.

# Apprentissage multi-instance (MIL)

- En apprentissage automatique, l'apprentissage multi-instance (MIL) est un type d'apprentissage supervisé ;
- L'apprentissage multi-instance (MIL) utilisé dans cet article est basé sur un classificateur supervisé standard basé sur des machines à vecteurs de support ;
- Ce classificateur utilise la perte de charnière (hinge loss) comme fonction de perte pour les vidéos d'apprentissage du classificateur ;
- Dans cette approche, les vidéos normales et anormales sont considérées comme des sacs et les segments vidéo comme des instances dans l'apprentissage par instances multiples (MIL).

$$\min_{\mathbf{w}} \left[ \frac{1}{z} \sum_{j=1}^z \max(0, 1 - Y_{\mathcal{B}_j} (\max_{i \in \mathcal{B}_j} (\mathbf{w} \cdot \phi(x_i)) - b)) \right] + \|\mathbf{w}\|^2$$



# Le modèle de classement MIL profond :

- Si les annotations au niveau des segments sont connues, alors :

$$f(\mathcal{V}_a) > f(\mathcal{V}_n)$$

- Tri par apprentissage d'instances multiples (MIL) pour la détection d'anomalies :

$$\max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) > \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i)$$

- Au lieu d'imposer un classement pour chaque instance du sac, la méthode impose un classement uniquement pour **les deux instances ayant les scores anormaux les plus élevés** dans les sacs positif et négatif, respectivement.
  - Score d'anomalie le plus élevé dans le sac positif → instances vraisemblablement positives
  - Score d'anomalie le plus élevé dans le sac négatif → ressemble le plus aux instances anormales

# Le modèle de classement MIL profond :

- Perte de classement dans la formulation de la perte de charnière :

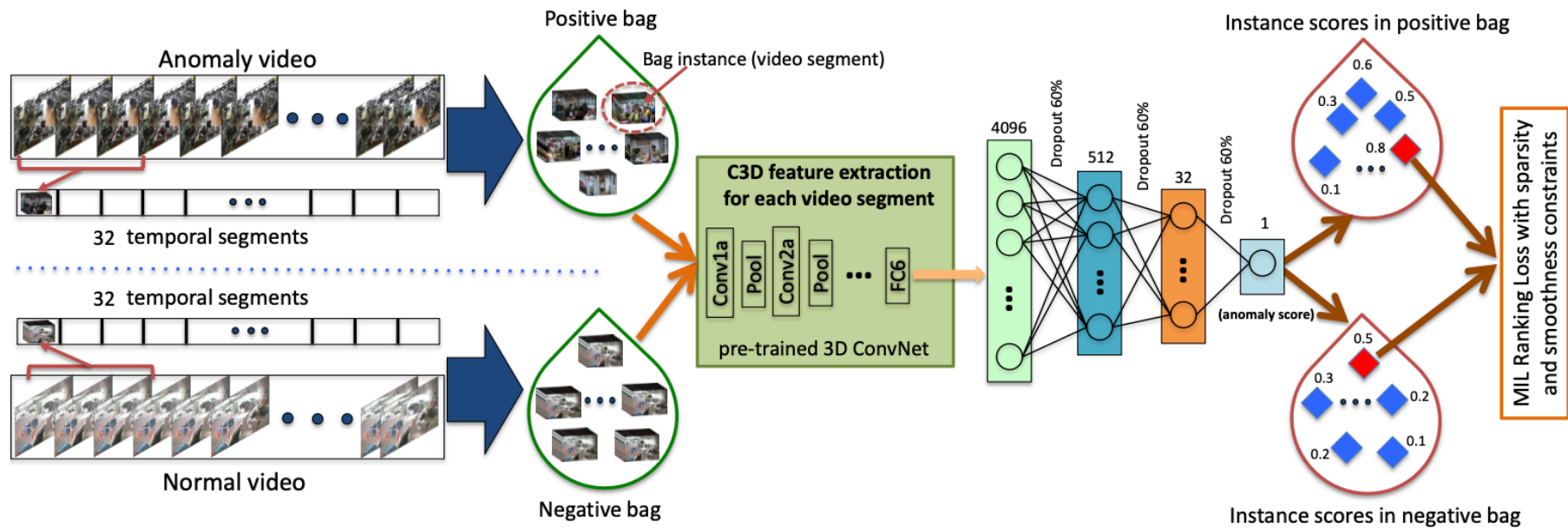
$$l(\mathcal{B}_a, \mathcal{B}_n) = \max(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i))$$

- Contraintes de régularité (1) et d'éparpillement (2) :

$$l(\mathcal{B}_a, \mathcal{B}_n) = \max(0, 1 - \max_{i \in \mathcal{B}_a} f(\mathcal{V}_a^i) + \max_{i \in \mathcal{B}_n} f(\mathcal{V}_n^i))$$

$$+ \lambda_1 \overbrace{\sum_i^{(n-1)} (f(\mathcal{V}_a^i) - f(\mathcal{V}_a^{i+1}))^2}^{\textcircled{1}} + \lambda_2 \overbrace{\sum_i^n f(\mathcal{V}_a^i)}^{\textcircled{2}}$$

# Cadre de détection des anomalies



# **Un ensemble de données pour la détection des anomalies**

## **Limites des ensembles de données précédents :**

- Anomalies simples (mauvaise orientation, course, etc.) ;
- Anomalies mises en scène et irréalistes ;
- La vidéo est tournée à un seul endroit.

# **Un ensemble de données pour la détection des anomalies**

## **Cet ensemble de données de détection d'anomalies :**

- 1 900 128 heures de vidéos de surveillance réelles ;
- 15 fois plus de vidéos que les ensembles de données existants ;
- 13 anomalies réelles.



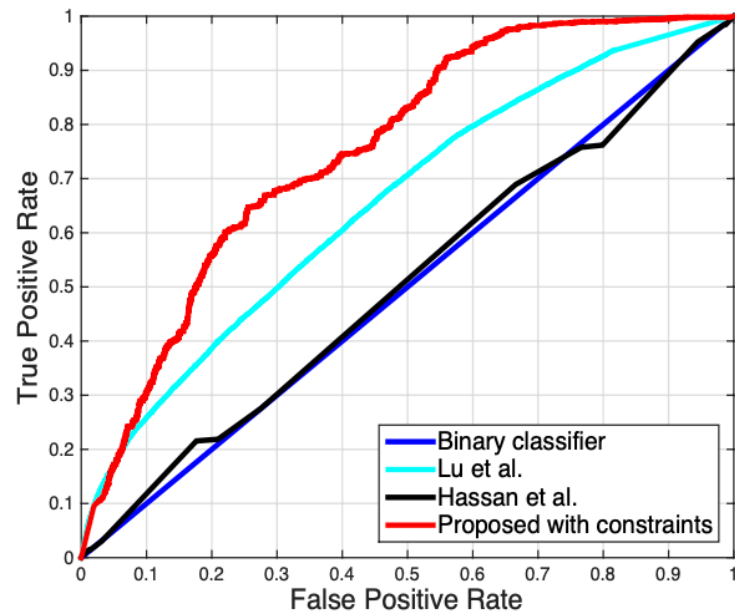
# Expériences et résultats

## Expériences

- Ensemble d'entraînement : 800 vidéos normales et 810 vidéos anormales ;
- Ensemble de tests : 150 vidéos normales restantes et 140 anomalies.

# Expériences et résultats

## Résultats de la détection des anomalies



Method	AUC
Binary classifier	50.0
Hasan <i>et al.</i> [18]	50.6
Lu <i>et al.</i> [28]	65.51
Proposed w/o constraints	74.44
<b>Proposed w constraints</b>	<b>75.41</b>

# En résumé

- Cette recherche développe une méthode de détection d'anomalies utilisant un cadre profond de tri multi-instances avec des données faiblement étiquetées ;
- Un nouvel ensemble de données d'anomalies à grande échelle est introduit, qui consiste en diverses anomalies dans le monde réel ;
- La performance de la méthode de détection d'anomalies proposée dans cette étude est significativement meilleure que la méthode de référence ;
- La modification de l'ensemble de données de l'étude fournit également une référence stimulante pour l'identification des activités anormales et offre des possibilités supplémentaires pour les travaux futurs.